

PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau



INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 : C12N 15/09, 15/12, 15/33, 15/64		A1	(11) International Publication Number: WO 96/09378 (43) International Publication Date: 28 March 1996 (28.03.96)
(21) International Application Number: PCT/US95/11511 (22) International Filing Date: 8 September 1995 (08.09.95)		(81) Designated States: AU, BG, BR, BY, CA, CN, CZ, FI, HU, JP, KR, MX, NO, NZ, PL, RO, RU, SG, SI, UA, European patent (AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE). Published <i>With international search report.</i>	
(30) Priority Data: 08/324,243 19 September 1994 (19.09.94) US			
(71) Applicant: THE GENERAL HOSPITAL CORPORATION [US/US]; 55 Fruit Street, Boston, MA 02114 (US).			
(72) Inventor: SEED, Brian; Apartment 5J, Nine Hawthorne Place, Boston, MA 02114 (US).			
(74) Agent: LECH, Karen, F.; Fish & Richardson P.C., 225 Franklin Street, Boston, MA 02114 (US).			
(54) Title: OVEREXPRESSION OF MAMMALIAN AND VIRAL PROTEINS			
(57) Abstract <p>The invention features a synthetic gene encoding a protein normally expressed in mammalian cells wherein at least one non-preferred or less preferred codon in the natural gene encoding the mammalian protein has been replaced by a preferred codon encoding the same amino acid.</p>			

- 55 -

1. A synthetic gene encoding a protein normally expressed in mammalian cells wherein at least one non-preferred or less preferred codon in the natural gene encoding said mammalian protein has been replaced by a preferred codon encoding the same amino acid.
5
2. The synthetic gene of claim 1 wherein said synthetic gene is capable of expressing said mammalian protein at a level which is at least 110% of that expressed by said natural gene in an in vitro mammalian cell culture system under identical conditions.
10
3. The synthetic gene of claim 1 wherein said synthetic gene is capable of expressing said mammalian protein at a level which is at least 150% of that expressed by said natural gene in an in vitro cell culture system under identical conditions.
15
4. The synthetic gene of claim 1 wherein said synthetic gene is capable of expressing said mammalian protein at a level which is at least 200% of that expressed by said natural gene in an in vitro cell culture system under identical conditions.
20
5. The synthetic gene of claim 1 wherein said synthetic gene is capable of expressing said mammalian protein at a level which is at least 500% of that expressed by said natural gene in an in vitro cell culture system under identical conditions.
25
6. The synthetic gene of claim 1 wherein said synthetic gene is capable of expressing said mammalian protein at a level which is at least ten times that expressed by said natural gene in an in vitro cell culture system under identical conditions.
30

- 56 -

7. The synthetic gene of claim 1 wherein at least 10% of the codons in said natural gene are non-preferred codons.

8. The synthetic gene of claim 1 wherein at least 5 50% of the codons in said natural gene are non-preferred codons.

9. The synthetic gene of claim 1 wherein at least 50% of the non-preferred codons and less preferred codons present in said natural gene have been replaced by 10 preferred codons.

10. The synthetic gene of claim 1 wherein at least 90% of the non-preferred codons and less preferred codons present in said natural gene have been replaced by preferred codons.

15 11. The synthetic gene of claim 1 wherein said protein is a retroviral or lentiviral protein.

12. The synthetic gene of claim 11 wherein said protein is an HIV protein.

20 13. The synthetic gene of claim 12 wherein said protein is selected from the group consisting of gag, pol, and env.

14. The synthetic gene of claim 13 wherein said protein is gp120 or gp160.

25 15. The synthetic gene of claim 1 wherein said protein is a human protein.

- 57 -

16. A method for preparing a synthetic gene encoding a protein normally expressed by mammalian cells, comprising identifying non-preferred and less-preferred codons in the natural gene encoding said protein and
5 replacing one or more of said non-preferred and less-preferred codons with a preferred codon encoding the same amino acid as the replaced codon.

Syngp120mn

1 CTCGAGATCC ATTGTGCTCT AAAGGAGATA CCCCCCAGA CCCCCCTCAC
 51 TCGGGTGCCT AGCTGCCCAAG CCTGAGGCAA GAGAAGGCAA GAAACCATGC
 101 CCATGGGTC TGTGCAACG CTGGCCACCT TGTACCTGCT CGGATGCTG
 151 CTGCTTCCG TGTAGCCAC CGAGAAGCTG TGGGTGACCG TGTACTACGG
 201 CGTGCCCTG TGTAGGAGG CGACCACAC CGTGTGCTG CGAGGGACG
 251 CCAAGGCTA CGACACCGAG CTGACACACG TGTGGCCAC CGAGGCTGC
 301 GTGCGACCG ACGGAAACCG CGAGGAGCTG GAGCTGTGA ACGTGACCGA
 351 GAACTTCAAC ATGTGGAAAGA ACACACATGTG CGACCGAGTG CATGAGGACA
 401 TGATCACTT GTGGGACCG AGCTGAAAGC CGTGTGAA CGTGACCCCG
 451 CTGTGCTGA CGTGTGACTG CACCGACCTG AGGAACACCA CGAACACCAA
 501 CAAAGGACCC GCGAACACCA ACAGGAAACG CGAGGGACCC ATCAAGGGCG
 551 CGGAGATGAA CAACTGCAGC TTCAACATCA CGACCGAGCT CGCGACACAG
 601 ATGGAGAAGG ATGACGCTT CGTGTGACCG CTGGATATCG TTAGCATCGA
 651 CAACGACAGC ACCACCTACC CGCTGATCTC CTGCAACACCG ACCGTGATCA
 701 CCCAGGCTG CGCCAAAGATC AGCTTGAGC CGATCCCCAT CGACTACTGC
 751 CGCCCGCCG CGTGTGCTG CGTGTGAGTGC AAACGAAAGA AGTTGAGGG
 801 CAAGGGCAGC TGTAAAGAAAG TTAGGACCCCT CGACTGACCC CACGGCATCC
 851 CGCCCGTGTG TGGCAGCCAG CTGTGCTGA AGGGCAGCT CGCGAGGAG
 901 GAGGTGTGAA TCCCGAGGCA GAACTTCAACG GACAAACCCG AGACCATCAT
 951 CGTGGACCTG AATGAGAGGG TGAGCATCAA CTGCAACCGT CGCAACTACA
 1001 ACAACGSCAA CGCCATCCAC ATCGGCCCCCG CGCGGCGCTT CTACACCCACC
 1051 AAGAACATCA TCGGCACCAT CGCCCAAGCC CACTGCAACCA TGTGTAGAGC
 1101 CGAGTGGAAC GACACCTGC CGCAGATCTG GAGCAAGCTG AAGGAGGAGT
 1151 CGAAGAACAA GACCATCGTG TTCAACCCAGA CGAGGCGGG CGACCCCCAG
 1201 ATCGTGTGATGC ACAGCTTCAA CTGGGGCGGC GAATTCTT ACTGCAACAC
 1251 CGGCCCCCTG TTCAACAGCA CGTGTGAAACCG CAAACACCCG TGGAAACAAACA
 1301 CGACCGGGAG CGAACACAAAT ATTACCTCC AGTGTGAGAT CGAGCAGATC
 1351 ATCAACATGTG CGCAGGAGGT CGCCCAAGCC ATGTACGCCC CGCCCATCGA
 1401 CGGCCAGATC CGGTGGAGCA CGAACATCAC CGTGTGCTG CTGACCCCCGG
 1451 ACGGGCGSCAA CGACACCGAC ACCAACGACA CGGAAATCTT CGGGGGGGC

FIG 1
(SHEET 1 OF 4)

2/12

1501 GCGGCGACCA TCGCCGACAA CTGGAGATCT GAGCTTACA AGTACAAAGGT
1551 CCTGACCGATC GAGCCCGCTGG CGCTGGGGCC CACCAAGGCC AAGCCGGGCG
1601 TGGTCCAGCG CAGAAGGCCTT AAAGGGCC CC (SEQ ID NO:34)

FIG 1
(SHEET 2 OF 4)

Syngp160mn

1 AUCGAGAAGC TGTGGGTGAC CTTTCTACTAC CGCGTACCGG TGTGGAAAGGA
 31 GCGCCACCAACG AGCTTGTTCAT GCGCCGACCGA CGCGAAGGGG TACGACACCS
 61 ATGTTGACAAA CTTGTTGGCC ACCGGAGGTTT GCGTGCCTCAC CGACCCCGAAC
 91 CGCGAGGAGG TTGAGCTTGT GAACTTACCC GAGAACTTCA ACATGTGAA
 121 GAACAAACATG CTGGAGGAGA TGTATGAGGA GATCATCAGC CTGTGGGACC
 151 AGAGCTGAA GCGCTTACGAGA AGCTGACCC CGCTTGCTGT GACCTTCAC
 181 TCGACCGATT TTAGGAGACG GACCAACCCG AACAAACAGCA CGCGCGGAA
 211 GAAACGGAGAC AGCGAGGCGA CGATCAAGGG CGCGCGAGATG AAUAACCTGCA
 241 GTTTCACAT CACCAACCCG AGCGCGGACCA AGATCGAGAA CGAGTACGCC
 271 CCTGTCTACA AGCTGAGTAT CGTGGACGTC GACAAACGAA CGACCGACCA
 301 CGCGCTTACG TGTGCGAAACG CGACCGTGT GAGCGAGGGG CGCGCGAAAGA
 331 CGACGTTGCA GCGCTTGGGG AGCGAGTACG CGCGCGGGCG CGCGCTTGGC
 361 ATCGTGTACGT GCGACGACAA GAACTTCAAGG CGCGAAGGGG CGTGCAGAGAA
 391 CGTGCAGGAGG :TGTGAG.GCA CGCGCGGCGT CGCGCGCGT CGTGCAGGACCG
 421 AGCTCTGTGAT GAAAGCGGAGG CGCGCGCGT GAGAGGTGTG TATTCGAGCG
 451 GAGAACTTGA CGCGCGACCG CGAGACCGAT ATCGTGCAGC TGTATGAGAG
 481 CGTGCAGGAGG AACTGACCGC CGCGCGACTA CGAAGACGCG AGCGCGACCG
 511 ACATCGCGCC CGCGCGCGCG CGTGTACGCA CGAAGAACAT CGCGCGACCG
 541 ATCGCGCGCG CGCGCGCGCG CGATGTGAGA CGCGAAGTGGG AGCGACCGCGT
 571 CGCGCGCGCGT GTGAGGAGCA CGTGCAGGAGG CGCGCGCGCG AGCGACCGCG
 601 CGTGCAGGAGG GAGGAGGAGG CGCGCGCGCG AGATCGTGTG CGACGCGTTC
 631 AACTGCGCGT GCGAATTGTT CGTGCAGGAGG AGCGACCGCG CGTGCAGGAG
 661 GACGTGGACG GCGAACAAACA CGCGCGACAA CGCGCGCGCG CGACGACAAACA
 691 ATATTACCGT CGCGCGCGCG AGCGACCGAGA .CGATCGACAT CGCGCGCGAG
 721 CGCGCGCGCG CGATGTGAGC CGCGCGCGCGT GAGCGAGGAGA CGCGCGCGCG
 751 CGACGACATG AGCGCGTGTG CGTGCAGGAGG CGCGCGCGCG AGCGACCGCG
 781 ACACCGCGCG CAACCGCGCG CGCGCGCGCG CGCGCGCGCG CGATCGACCG
 811 AGATCGAGAT CGTGCAGGAGT GAACTGACGAG AGCGACCGCG CGTGCAGGAG
 841 CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG

4/12

1451 CGGCCCGCAT CGCCCGCGT TCTCTGGCT TCTTGGGGC CGCCCGCGC
1501 ACCATGGGGG CGCCCGAGGT GACCCCTGACC CTGAGGGCC CGCTGGCTCT
1551 GAGCGGCATC GTGCAGCAGC AGAACAAACCT CCTCCGGGGC ATCGAGGGCC
1601 AGCAGGATAT GTCAGGCTG ACCCTTGCG CGATCAAGCA GCTCGAGGGC
1651 CGCTGGCTCG CGCTGGAGGCT TACCTGAAG GACCGAGGAGC TCTTGGCGCT
1701 CTGGGGCTCG TCGGCGAAGC TGATCTGACG CACFAACGTA CGCTGGAAACG
1751 CCTCTGGAG CAACAAGAGC CTGGAGGAGA TCTGGAAACAA CATGACCTCG
1801 ATCGACTGGG AGCCCGAGAT CGATAACTAC ACCAGCGCTA TCTACAGGCT
1851 CCTGGAGAGG AGCCAGACCG ACCAGGAGAA GAACGAGCAG GAGCTCTGG
1901 ACCTGGACAA CTGGGGAGG CTCTGGAACT CGCTGGACAT CACCAACTCG
1951 CTCTGGTACA TCAAATATT CATCATGATT CTGGGGCGC TCTTGGCGCT
2001 CGCCATCGT TCTGGCGCTCG TCTGGATCTG CGCCCGCGT CGCCCGCGCT
2051 AGACGGGGCTG GAGCTGGAG ACCGGGGCGG CGCTGGCGCG CGGGGGGGAC
2101 CGCCCGCGCG CGATCGAGGA CGAGGGCGGC GAGCGCGACC CGCACACCG
2151 CGCGAGGCTG GTGCGACCGT TCTGGCGAT CGCTGGCGT CGCTGGCGCA
2201 CGCTGGCTCG GTCAGGCTAC CGACACCGCG ACCCTGGCTCT GATGGGGCGC
2251 CGCATCTGG AACTCTTAGG CGCCCGCGCG TGGGAGGTGC TGAAGTACTG
2301 CTGGAAACCTG CGCCAGTATT CGAGCCAGGA CGCTGGCTCG AGCCCGCTGA
2351 CGCTGGCTGA CGCCACCGCG ACCGGGGCGT CGAGGGCGAC CGCCCGCGT
2401 ATCGAGGCTUC TCGAGAGGGC CGGGAGGGCG ACCCTGACCA CGCCCGCGCG
2451 CGATGGCGCG CGCTGGAGGA CGCCCGCGT 3 (SEQ ID NO: 3)

FIG. 1
(SHEET 4 OF 4)

5/12

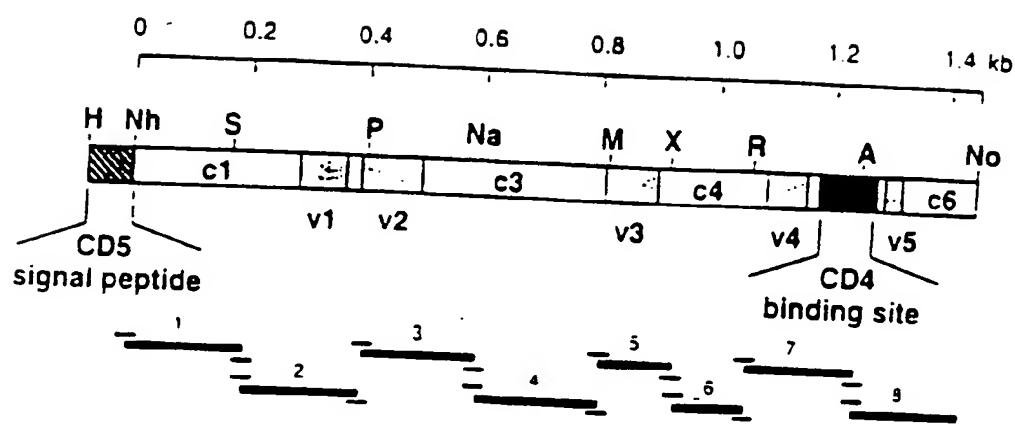


FIGURE 2

6/12

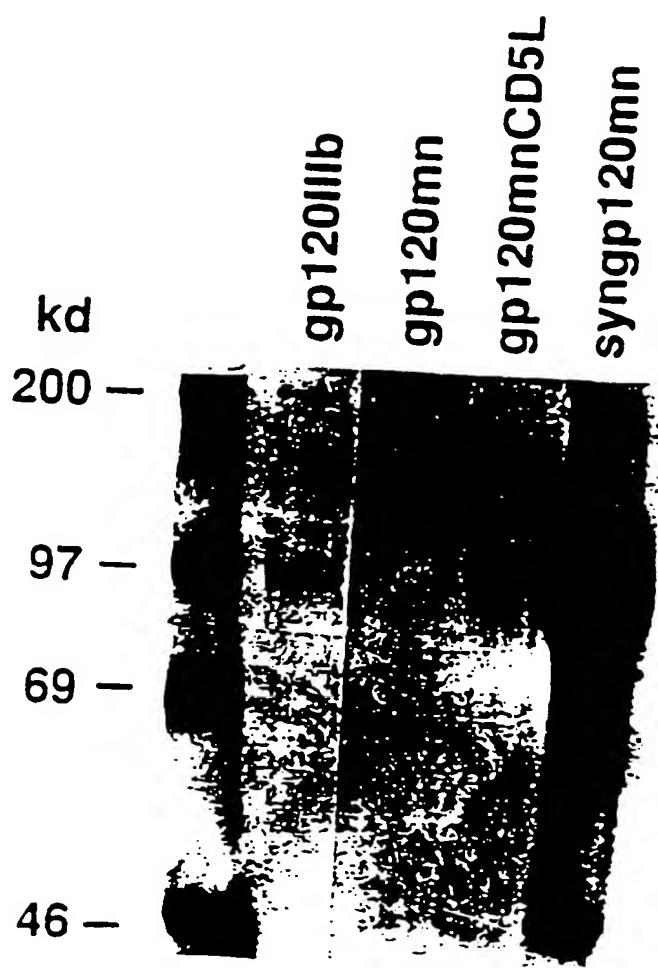


FIGURE 3

7/12

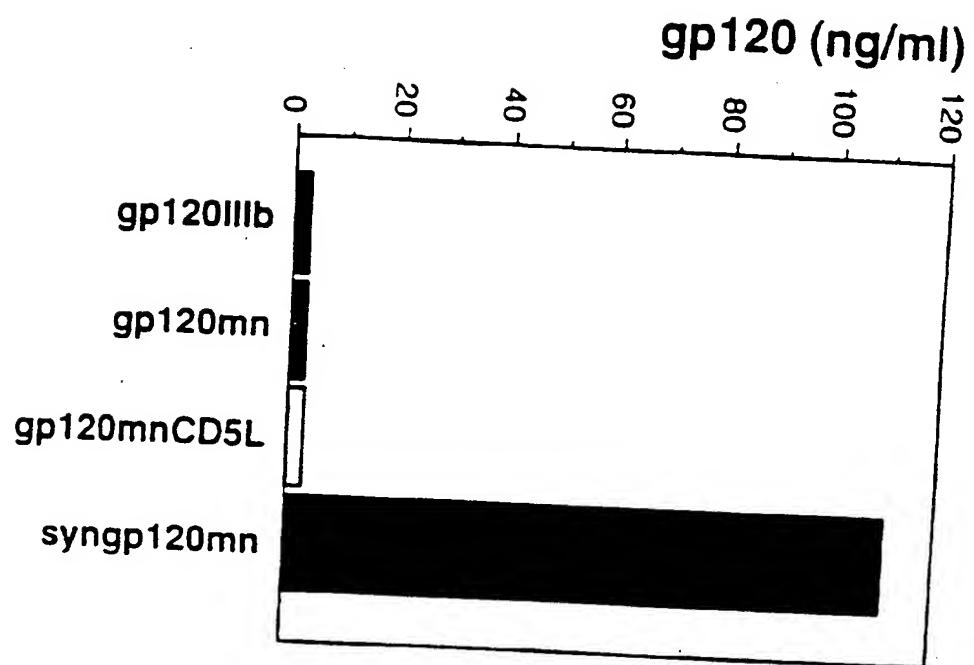


FIGURE 4

8/12

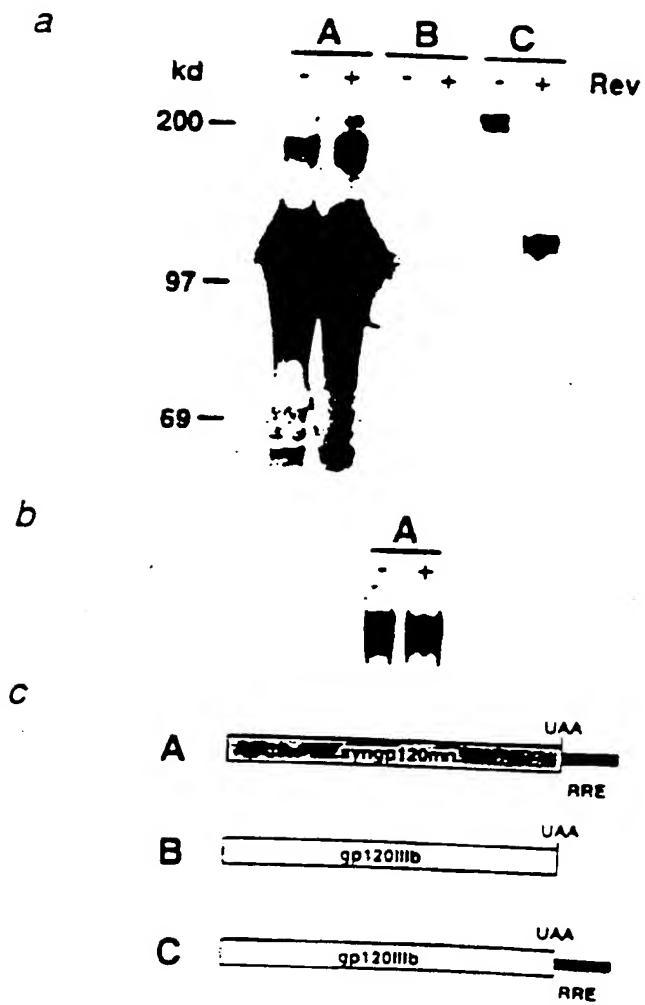


FIGURE 5

9 / 12

FIGURE 6

10/12

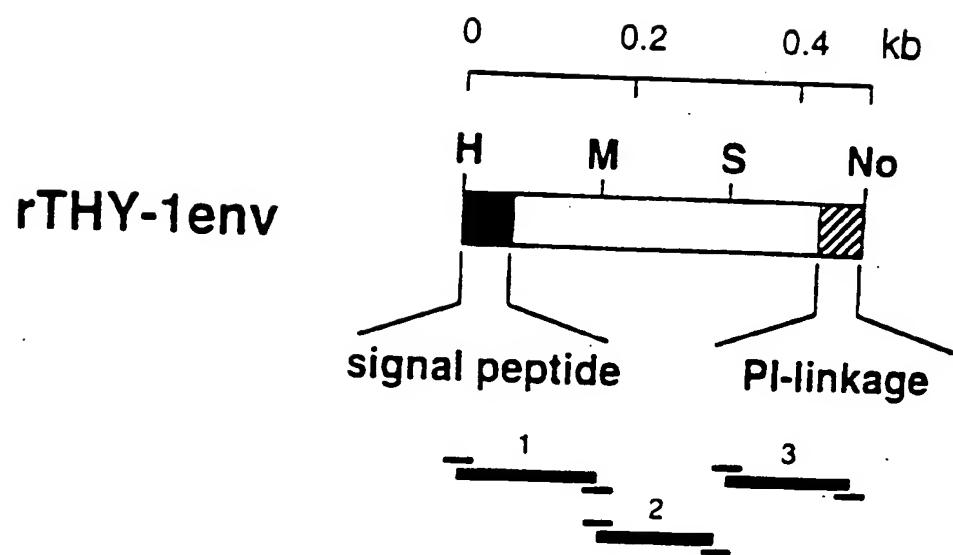


FIGURE 7

11/12

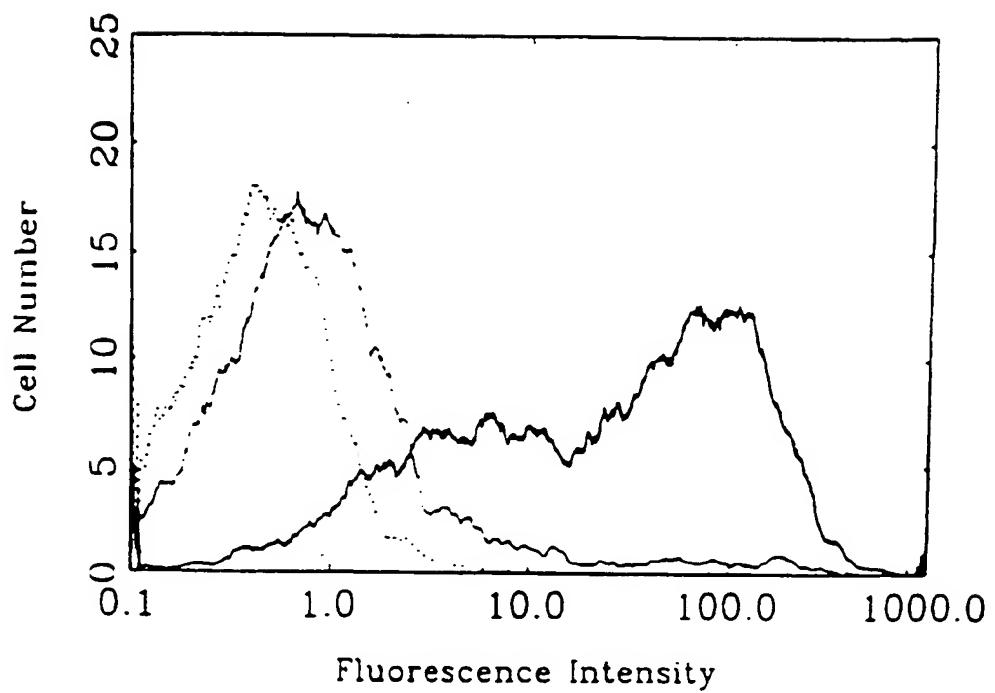


FIGURE 8

12/12

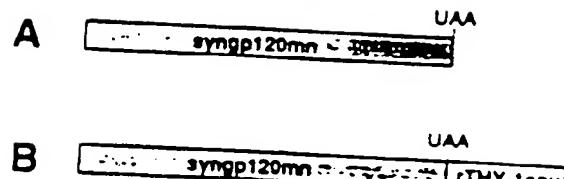
a*b*

FIGURE 9

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US95/11511

A. CLASSIFICATION OF SUBJECT MATTER

IPC(6) : Please See Extra Sheet.

US CL : Please See Extra Sheet.

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : Please See Extra Sheet.

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched
Please See Extra Sheet.

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)
Please See Extra Sheet.

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	US, A, 5,270,171 (CERCEK ET AL.) 14 December 1993, see column 34, lines 32-48.	1-16
Y	Nucleic Acids Research, Volume 18, Number 4, issued 1990, McCarrey, "Molecular evolution of the human Pfk-2 retroposon", pages 949-955, see entire document.	1-16
Y	Japanese Journal of Cancer Research, Volume 80, issued March 1989, Kamiya et al., "Transformation of NIH3T3 Cells with Synthetic c-Ha-ras Genes", pages 200-203, see entire document.	1-16

Further documents are listed in the continuation of Box C.

See patent family annex.

* Special categories of cited documents.	"T"	later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"A" document defining the general state of the art which is not considered to be of particular relevance	"X"	document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"E" earlier document published on or after the international filing date	"Y"	document of particular relevance; the claimed invention would be considered to involve an inventive step when the document is taken alone
"I" document which may throw doubt on priority claim(s) or which is used to establish the publication date of another citation or other special reason (as specified)	"Z"	document of particular relevance; the claimed invention would be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"O" document referring to an oral disclosure, use, exhibition or other means	"A"	document member of the same patent family
"P" document published prior to the international filing date but later than the priority date claimed		

Date of the actual completion of the international search

24 OCTOBER 1995

Date of mailing of the international search report

03 NOV 1995

Name and mailing address of the ISA/US
Commissioner of Patents and Trademarks
Box PCT
Washington, D.C. 20231

Faximile No. (703) 305-3230

Authorized officer

JAMES KETTER

Telephone No. (703) 308-0196

INTERNATIONAL SEARCH REPORT

In. national application No.
PCT/US95/11511

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	Nucleic Acids Research, Volume 16, Number 17, issued 1988, Sharp et al., "Codon usage patterns in <i>Escherichia coli</i> , <i>Bacillus subtilis</i> , <i>Saccharomyces cerevisiae</i> , <i>Schizosaccharomyces pombe</i> , <i>Drosophila melanogaster</i> and <i>Homo sapiens</i> : a review of the considerable within-species diversity", pages 8207-8211, see entire document.	1-16
Y	Proceedings of the National Academy of Sciences USA, Volume 83, issued November 1986, Newgard et al., "Sequence analysis of the cDNA encoding human liver glycogen phosphorylase reveals tissue-specific codon usage", pages 8132-8136, see entire document.	1-16
Y	Gene, Volume 46, issued 1986, Coulombe et al., "Expression of a synthetic human interferon- α , gene with modified nucleotide sequence in mammalian cells", pages 89-95, see entire document.	1-16

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US95/11511

A. CLASSIFICATION OF SUBJECT MATTER:
IPC (6):

C12N 15/09, 15/12, 15/33, 15/64

A. CLASSIFICATION OF SUBJECT MATTER:
US CL :

536/23.5, 23.72; 435/172.3

B. FIELDS SEARCHED

Minimum documentation searched

Classification System: U.S.

536/23.5, 23.72; 435/172.3

B. FIELDS SEARCHED

Documentation other than minimum documentation that are included in the fields searched:

NONE

B. FIELDS SEARCHED

Electronic data bases consulted (Name of data base and where practicable terms used):

APS, MEDLINE EXPRESS